



Best Available Copy

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

*[Handwritten signature]*

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/812.000	03/30/2004	Tetsuji Nakagawa	31869-202101	2668

26694 7590 01/25/2008  
VENABLE LLP  
P.O. BOX 34385  
WASHINGTON, DC 20043-9998

EXAMINER

GODDARD, DOUGLAS

ART UNIT	PAPER NUMBER
----------	--------------

2626

MAIL DATE	DELIVERY MODE
-----------	---------------

01/25/2008

PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

<b>Office Action Summary</b>	<b>Application No.</b>	<b>Applicant(s)</b>	
	10/812,000	NAKAGAWA, TETSUJI	
	<b>Examiner</b>	<b>Art Unit</b>	
	Douglas C. Godbold	2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

#### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

#### Status

- 1) ☐ Responsive to communication(s) filed on 31 October 2007.
- 2a) ☒ This action is **FINAL**.                      2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

#### Disposition of Claims

- 4) ☒ Claim(s) 1-20 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-20 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

#### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

#### Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \* c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
  - ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
  - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- \* See the attached detailed Office action for a list of the certified copies not received.

#### Attachment(s)

- |  |   |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)          | 4) <input type="checkbox"/> Interview Summary (PTO-413)           |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____                                      |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)          | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date. _____   | 6) <input type="checkbox"/> Other: _____                          |

### **DETAILED ACTION**

1. This office action is in response to correspondence filed October 31, 2007 in reference to application 10/812,000. Claims 1-20 are pending in the application and have been examined.

#### ***Response to Amendment***

2. The amendments to the claims filed October 31, 2007 have been accepted and considered in this office action. Claims 1, 2, 4, 8, 11, 13, 14, 16-18 and 20 have been amended.

#### ***Response to Arguments***

3. Applicant's arguments with respect to claims 1- 20 have been considered but are moot in view of the new ground(s) of rejection.

#### ***Claim Rejections - 35 USC § 103***

1. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

2. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

3. Claims 1, 13, and 20 rejected under 35 U.S.C. 103(a) as being unpatentable over Nagata (A Stochastic Japanese Morphological Analyzer Using a Forward-DP Backward-A N-Best Search Algorithm) in view of Aires (Combining Multiple Classifiers to Improve Part of Speech Tagging.)

4. Claim 1, 13, and 20 are rejected under 35 U.S.C. 102(b) as being anticipated by Nagata (A Stochastic Japanese Morphological Analyzer Using a Forward-DP Backward-A N-Best Search Algorithm).

5. Consider claim 1, Nagata teaches a morphological analyzer comprising (Japanese morphological analyzer, abstract):

a hypothesis generator for applying a prescribed method of morphological analysis to a text and generating one or more hypotheses as candidate results of the morphological analysis, each hypothesis being a word string with part-of-speech tags, the part-of-speech tags including form information for parts of speech having forms (We propose a novel search strategy for getting the N best morphological analysis hypotheses for the input sentence; page 201, column 1, line 35. We used tile simple tri-POS model as the tagging model for Japanese; page 201, column 1, line 30.);

a model storage facility storing information for a plurality of part-of-speech n-gram models, at least one of the part-of-speech n-gram models including information about the forms of the parts of speech (We used tile simple tri-POS model as the tagging model for Japanese; page 201, column 1, line 30. Tagging model also described in detail in section 2. It is inherent that the model must be stored in a memory in order to enable the analyzer.);

a probability calculator for finding a probability that each said hypothesis will appear in a large corpus of text by using a weighted combination of the information for the part-of-speech n-gram models stored in the model storage facility (equations 3, 4 and 5 described in section 201 estimates the probabilities of the tagging with relation to the relative frequencies of the corresponding events.); and

a solution finder for finding a solution among said hypotheses, based on the probabilities generated by the probability calculator (Once word hypotheses for unknown words are generated, the proposed N-best algorithm will find tile most likely word segmentation and part of speech assignment taking into account the entire sentence; page 204, column 1, line 22.)

Nagata does not specifically teach at least two of the part-of-speech n-gram models being based on mutually different types of morphological information.

In the same field of POS tagging, Aires teaches at least two of the part-of-speech n-gram models being based on mutually different types of morphological information (four different classifiers were used, including a Unigram and an n-gram tagger; page 2 paragraph 1. The unigram and n-gram are two different n-gram models that are based

on different morphological information as the unigram considers just one word/morpheme at a time and the n-gram considers the context of n words/morphemes.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the morphological analyzer of Nagata with the multiple part of speech models as taught by Aires in order to obtain a higher accuracy in tagging, (abstract Aires).

6. Consider claim 13, A method of morphological analysis comprising (Japanese morphological analyzer, abstract):

applying a prescribed method of morphological analysis to a text and generating one or more hypotheses as candidate results of the morphological analysis, each hypothesis being a word string with part-of-speech tags, the part-of-speech tags including form information for parts of speech having forms (We propose a novel search strategy for getting the N best morphological analysis hypotheses for the input sentence; page 201, column 1, line 35. We used tile simple tri-POS model as the tagging model for Japanese; page 201, column 1, line 30.);

calculating probabilities that each said hypothesis will appear in a large corpus of text by using a weighted combination of a plurality of part-of-speech n-gram models, at least one of the part-of-speech n-gram models including information about forms of parts of speech (equations 3, 4 and 5 described in section 201 estimates the

probabilities of the tagging with relation to the relative frequencies of the corresponding events.); and

finding a solution among said hypotheses, based on said probabilities (Once word hypotheses for unknown words are generated, the proposed N-best algorithm will find the most likely word segmentation and part of speech assignment taking into account the entire sentence; page 204, column 1, line 22.).

Nagata does not specifically teach at least two of the part-of-speech n-gram models being based on mutually different types of morphological information.

In the same field of POS tagging, Aires teaches at least two of the part-of-speech n-gram models being based on mutually different types of morphological information (four different classifiers were used, including a Unigram and an n-gram tagger; page 2 paragraph 1. The unigram and n-gram are two different n-gram models that are based on different morphological information as the unigram considers just one word/morpheme at a time and the n-gram considers the context of n words/morphemes.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the morphological analyzer of Nagata with the multiple part of speech models as taught by Aires in order to obtain a higher accuracy in tagging, (abstract Aires).

7. Consider claim 20, Nagata teaches a machine-readable tangible medium storing a program comprising instructions that can be executed by a computing device to carry

out morphological analysis by the method of claim 13 (Nagata teaches a morphological analyzer; abstract. In order to operate this analyzer it is inherent that there is in fact a storage medium containing the instructions in order to operate it. ).

8. Claims 2, 3, 14, and 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nagata and Aires as applied to claims 1 and 13 above and further in view of Haruno et al. (Mistake-Driven Mixture of Hierarchical Tag Context Trees).

9. Consider claim 2, Nagata and Aires teaches the morphological analyzer of claim 1, but does not specifically teach wherein said at least one of the part-of-speech n-gram models including information about forms of parts of speech is a hierarchical part-of-speech n-gram model that treats parts of speech and their forms at different hierarchical levels.

In the same field of tagging, Haruno teaches a part-of-speech n-gram model including information about forms of parts of speech is a hierarchical part-of-speech n-gram model that treats parts of speech and their forms at different hierarchical levels (the tags are representing in a hierarchical tag context tree, abstract. Figure 1 shows parts of speech and their forms on different hierarchical levels.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the hierarchical model of Haruno with the morphological system of Nagata and Aires in order to use a method of tagging that outperforms both hand-crafted and conventional statistical methods, abstract Aires.



Consider claim 3, Haruno suggests the morphological analyzer of claim 2, wherein the hierarchical part-of-speech n-gram model calculates a product of a conditional probability  $P(w_i | t_i)$  of occurrence of a word  $w_i$  given its part of speech  $t_i$ , a conditional probability  $P(tiform | tipos)$  of occurrence of the part of speech  $tipos$  of said word  $w_i$  in a form  $tiform$  shown by said word  $w_i$ , and a conditional probability  $P(tipos | t_{i-N+1} \dots t_i)$  of occurrence of the part of speech  $tipos$  of said word  $w_i$  following a part-of-speech tag string  $t_{i-N+1} \dots t_i$  indicating parts of speech of  $N - 1$  preceding words, where  $N$  is a positive integer (Page 231, equation 1 shows multiplying  $P(W_i | t_i)$  with  $P(t_i | t_{1,i-1})$ ). It would be obvious to one of ordinary skill in the art when considering separate forms of a word to add the conditional probability of  $P(tiform | tipos)$  to the product in order to provide The probability of the further subdivisions of the part of speeches.

10. Consider claim 14, Nagata teaches method of claim 13, but does not specifically teach wherein said at least one of the part-of-speech n-gram models including information about forms of parts of speech is a hierarchical part-of-speech n-gram model.

In the same field of tagging, Haruno teaches a part-of-speech n-gram model including information about forms of parts of speech is a hierarchical part-of-speech n-gram model that treats parts of speech and their forms at different hierarchical levels (the tags are representing in a hierarchical tag context tree, abstract. Figure 1 shows parts of speech and their forms on different hierarchical levels.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to combine the hierarchical model of Haruno with the morphological system of Nagata and Aires in order to use a method of tagging that outperforms both hand-crafted and conventional statistical methods, abstract Aires.

Consider claim 15, Nagata teaches the method of claim 14, wherein the hierarchical part-of-speech n-gram model calculates a product of a conditional probability  $P(w_i | t_i)$  of occurrence of a word  $w_i$  given its part of speech  $t_i$ , a conditional probability  $P(t_i | \text{form } i)$  of occurrence of the part of speech  $t_i$  of said word  $w_i$  in a form  $t_i$  shown by said word  $w_i$ , and a conditional probability  $P(t_i | \text{Pos } i | t_{i-N+1} \dots t_{i-1})$  of occurrence of the part of speech  $t_i$  of said word  $w_i$  following a part-of-speech tag string  $t_{i-N+1} \dots t_{i-1}$  indicating parts of speech of  $N - 1$  preceding words, where  $N$  is a positive integer (Page 231, equation 1 shows multiplying  $P(W_i | t_i)$  with  $P(t_i | t_1, i-1)$ ). It would be obvious to one of ordinary skill in the art when considering separate forms of a word to add the conditional probability of  $P(t_i | \text{form } i)$  to the product in order to provide The probability of the further subdivisions of the part of speeches.

11. Claims 4-7, and 16 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nagata in view of Aires as applied to claims 1 and 13 above and further in view of Lee et al (Lexicalized HMM's for Part-of-Speech Tagging).

12. Consider claim 4, Nagata in view of Aires teaches the morphological analyzer of claim 1, but does not specifically teach wherein at least one of the part-of-speech n-gram models is a lexicalized part-of-speech n-gram model in which all words are lexicalized.

In the same field of part of speech tagging, lee teaches at least one of the part-of-speech n-gram models is a lexicalized part-of-speech n-gram model (This paper discusses uniformly lexicalizing HMM models, that is lexicalizing all HMM models, instead of doing so selectively; abstract and introduction paragraphs 2 and 3).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to used lexicalized HMMs as taught by Lee with the analyzer of Nagata and Aires in order to improve tagging accuracy; Lee abstract.

13. Consider claim 5, Lee suggests the morphological analyzer of claim 4, wherein the lexicalized part-of-speech n-gram model calculates a product of a conditional probability  $P(w_i | t_i)$  of occurrence of a word  $w_i$  given its part of speech  $t_i$  and a conditional probability  $P(t_i | w_{i-N+1}t_{i-N+1} \dots w_{i-1}t_{i-1})$  of occurrence of the part of speech  $t_i$  of said word  $w_i$  following  $N - 1$  words  $w_{i-N+1} w_{i-1}$  having respective parts of speech  $t_{i-N+1} \dots t_{i-1}$ , where  $N$  is a positive integer (Equation 4 and Equation 5 can be manipulated to create the production of  $P(w_i | t_i)$  and  $P(t_i | w_{i-N+1}t_{i-N+1} \dots w_{i-1}t_{i-1})$ . As equation 5 is a derivative of equation 4 for lexicalized HMM's). ).

14. Consider claim 6, Lee teaches the morphological analyzer of claim 4, wherein the lexicalized part-of-speech n-gram model calculates a conditional probability  $P(w_i | t_{i-N+1} \dots t_{i-1})$  of occurrence of a word  $w_i$  having a part of speech  $t_i$  following a string of  $N - 1$  parts of speech  $t_{i-N+1} \dots t_{i-1}$ , where  $N$  is a positive integer (The probability of the current tag on the current word depends of the previous  $k$  tags; section 3.1 Therefore the equation of claim 6 is in obvious computation to make.).

15. Consider claim 7, Lee teaches the morphological analyzer of claim 4, wherein the lexicalized part-of-speech n-gram model calculates a conditional probability  $P(w_i | w_{i-N+1} t_{i-N+1} \dots w_{i-1} t_{i-1})$  of occurrence of a word  $w_i$  having a part of speech  $t_i$  following a string of  $N - 1$  words  $w_{i-N+1} \dots w_{i-1}$  having respective parts of speech  $t_{i-N+1} \dots t_{i-1}$ , where  $N$  is a positive integer (The probability of the current tag on the current word depends of the previous  $k$  tags and the  $K$  previous words; section 3.1 Therefore the equation of claim 7 is in obvious computation to make.)

16. Consider claim 16, Nagata and Aires teaches the method of claim 13, but does not specifically teach wherein at least one of the part-of-speech n-gram models is a lexicalized part-of-speech n-gram model in which all words are lexicalized.

In the same field of part of speech tagging, Lee teaches at least one of the part-of-speech n-gram models is a lexicalized part-of-speech n-gram model (This paper discusses uniformly lexicalizing HMM models, that is lexicalizing all HMM models, instead of doing so selectively; abstract and introduction paragraphs 2 and 3).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use lexicalized HMMs as taught by Lee with the analyzer of Nagata in order to improve tagging accuracy; Lee abstract.

17. Claims 8, 11, 17, and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nagata in view of Aires, as applied to claims 1 and 13 above, and further in view of Schone et al (Language-independent Induction of Part of Speech Class Labels Using only Language Universals).

18. Consider claim 8, Nagata and Aires teaches the morphological analyzer of claim 1, but does not specifically teach wherein at least one of the part-of-speech n-gram models stored in the model storage facility is a class part-of-speech n-gram model employing classes obtained by clustering.

In the same field of n-gram language modeling, Schone teaches at least one of the part-of-speech n-gram models stored in the model storage facility is a class part-of-speech n-gram model employing classes obtained by clustering (This algorithm derives clusters and part of speech tag; page 1 column 2 paragraph 2. ).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use class n-gram models as taught by Schone for parsing as taught by Nagata and Aires in order to provide a more accurate part of speech analysis.

19. Consider claim 11, Schone teaches the morphological analyzer of claim 8, wherein the class part-of-speech n-gram model is trained from both a part-of-speech tagged corpus and a part-of-speech untagged corpus, clustering parameters obtained from the part-of-speech untagged corpus being used to cluster morphemes in the part-of-speech tagged corpus (Section 2 describes the clustering, and the that the only previous knowledge is previously known grammar rules. No tagged corpus is used. ).

20. Consider claim 17, Nagata and Aires teaches the method of claim 13, but does not specifically teach wherein at least one of the part-of-speech n-gram models stored in the model storage facility is a class part-of-speech n-gram model employing classes obtained by clustering.

In the same field of n-gram language modeling, Schone teaches at least one of the part-of-speech n-gram models stored in the model storage facility is a class part-of-speech n-gram model employing classes obtained by clustering (This algorithm derives clusters and part of speech tag; page 1 column 2 paragraph 2. ).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use class n-gram models as taught by Schone for parsing as taught by Nagata and Aires in order to provide a more accurate part of speech analysis.

21. Consider claim 18, Schone teaches method of claim 17, teaches wherein the class part-of-speech n-gram model is trained from both a part-of-speech tagged corpus and a part-of-speech untagged corpus (Section 2 describes the clustering, and the that

the only previous knowledge is previously known grammar rules. No tagged corpus is used. ).

22. Claim 9-10 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nagata and Aires in view of Schone as applied to claim 8 above, and further in view of Lee et al.

23. Consider claim 9, Nagata, Aires and Schone teaches the morphological analyzer of claim 8, but does not teach specifically wherein the class part-of-speech n-gram model calculates a product of a conditional probability  $P(w_i | t_i)$  of occurrence of a word  $w_i$  given its part of speech  $t_i$  and a conditional probability  $P(t_i | C_{i-N+1} t_{i-N+1} \dots C_{i-1} t_{i-1})$  of occurrence of said part of speech  $t_i$  following a string of  $N - 1$  words assigned to respective classes  $C_{i-N+1} \dots C_{i-1}$  with respective parts of speech  $t_{i-N+1} \dots t_{i-1}$ , where  $N$  is a positive integer.

In the same field of tagging, Lee teaches lexical part-of-speech n-gram model calculates a product of a conditional probability  $P(w_i | t_i)$  of occurrence of a word  $w_i$  given its part of speech  $t_i$  and a conditional probability  $P(t_i | C_{i-N+1} t_{i-N+1} \dots C_{i-1} t_{i-1})$  of occurrence of said part of speech  $t_i$  following a string of  $N - 1$  words assigned to respective classes  $C_{i-N+1} \dots C_{i-1}$  with respective parts of speech  $t_{i-N+1} \dots t_{i-1}$ , where  $N$  is a positive integer (The probability of the current tag on the current word depends of the previous  $k$  tags and the  $K$  previous words; section 3.1 This can be combined with the

induction probabilities of Schone described in section 3. Therefore the equation of claim 9 is in obvious computation to make. This can be combined)

Therefore it would have been obvious to apply the same rules of probability equations for tagging as taught by Lee to the class tagger of Nagata and Schone and Aires in order to provide a more accurate tagging, taking fully into consideration the class of which the words belong.

24. Consider claim 10, Nagata, Aires and Schone and Aires teaches the morphological analyzer of claim 8, but does not specifically teach wherein the class part-of-speech n-gram model calculates a product of a conditional probability  $P(w_i | C_{i-N+1} t_{i-N+1} \dots C_{i-1} t_{i-1})$  of occurrence of a word  $w_i$  having a part of speech  $t_i$  following a string of  $N - 1$  words in respective classes  $C_{i-N+1} \dots C_{i-1}$  With respective parts of speech  $t_{i-N+1} \dots t_{i-1}$ , where  $N$  is a positive integer.

In the same field of tagging, Lee teaches wherein the lexical part-of-speech n-gram model calculates a product of a conditional probability  $P(w_i | C_{i-N+1} t_{i-N+1} \dots C_{i-1} t_{i-1})$  of occurrence of a word  $w_i$  having a part of speech  $t_i$  following a string of  $N - 1$  words in respective words  $C_{i-N+1} \dots C_{i-1}$  With respective parts of speech  $t_{i-N+1} \dots t_{i-1}$ , where  $N$  is a positive integer (The probability of the current tag on the current word depends of the previous  $k$  tags and the  $K$  previous words; section 3.1 This can be combined with the induction probabilities of Schone described in section 3. Therefore the equation of claim 10 is in obvious computation to make. This can be combined)



Therefore it would have been obvious to apply the same rules of probability equations for tagging as taught by Lee to the class tagger of Nagata and Schone and Aires in order to provide a more accurate tagging, taking fully into consideration the class of which the words belong.

25. Claims 12 and 19 is rejected under 35 U.S.C. 103(a) as being unpatentable over Nagata and Aires as applied to claims 1 and 13 above, and further in view of Siu (Variable N-Grams and Extensions for Conversational Speech Modeling).

26. Consider claim 12, Nagata and Aires teaches the morphological analyzer of claim 1, but does not specifically teach further comprising a weight calculation unit using a leave-one-out method to calculate weights of the part-of-speech n-gram models.

In the same field of part of speech tagging, Siu teaches a weight calculation unit using a leave-one-out method to calculate weights of the part-of-speech n-gram models (To evaluate the effect of node pruning and merging on unseen data, we use leave-one-out (LOO) likelihood to estimate the distributions used in the distance measures; column 69, column 2, section B.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the leave one out method as taught by Siu with the analyzer of Nagata and Aires in order to provide a more robust method of training.

27. Consider claim 19, Nagata and Aires teaches the method of claim 13, but does not specifically teach further comprising a weight calculation unit using a leave-one-out method to calculate weights of the part-of-speech n-gram models.

In the same field of part of speech tagging, Siu teaches a weight calculation unit using a leave-one-out method to calculate weights of the part-of-speech n-gram models (To evaluate the effect of node pruning and merging on unseen data, we use leave-one-out (LOO) likelihood to estimate the distributions used in the distance measures; column 69, column 2, section B.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use the leave one out method as taught by Siu with the analyzer of Nagata and Aires in order to provide a more robust method of training.

### ***Conclusion***

28. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any


extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Douglas C. Godbold whose telephone number is (571) 270-1451. The examiner can normally be reached on Monday-Thursday 7:00am-4:30pm Friday 7:00am-3:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571) 272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DCG

  
PATRICK EDOUARD  
SUPERVISORY PATENT EXAMINER